Future Enhancement:

1. Item Recommendation for Hero

2. Hero Picking Suggestion

3. Text Analysis within team mates

Predictive Analytics on Defence Of The Ancient 2 CSCI 680 Data Science

-Patlolla Akhil Reddy
Z1803493

what is DOTA?

- RPG / MOBA -
  Role Play Game /
  Multiplayer Online Battle Arena
- 2 Teams (5 Each)
- Fighting to destroy Enemies Building

What am I trying to predict?

which team wins the match
based on the features which
reflect current status of the game

Where did the data come from ?
   STEAM WEB API/ Academic Torrents
Size of the data ?

   around 780 GB

Sampling (Non-probability Sampling):
   Sort Whole data based on GLOBAL RANK
   780GB to 400 MB (48 Million Records)

Building the model:

Random Forest Technique considering the
features such as
['gold', 'gold_spent', 'gold_per_min',
'xp_per_min', 'kills', 'deaths', 'assists', 'denies',

'last_hits', 'stuns', 'hero_damage',
'hero_healing', 'tower_damage']

```
#
etrics.roc_auc_score(test_labels.values, test_probs[:,1])
```

.50338291745162056

```
#
rint(metrics.classification_report(test_labels.values, test_preds))
```

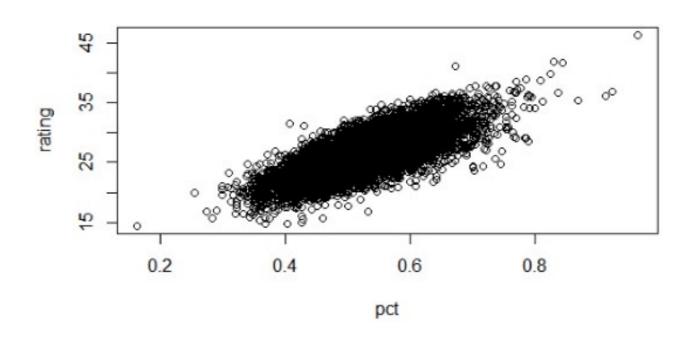|	precision	recall	f1-score	support
0	0.48	0.46	0.47	48139
1	0.52	0.55	0.53	51861
vg / total	0.50	0.50	0.50	100000

All Features:
ROC
Area Under Curve - 0.5
odd's of teams : almost 1 : 1

Win Prediction wrt Feature

Correlation: correlation of win percentage to all features true skill has highest value of **0.75192**

Building the model:

Random Forest Technique considering the
feature "True Skill"
True Skill Calculation:
GPM (gold per min), XPM (experience per min),
Randiant & Dire Tower Status wrt time,

It Worked !!

ROC -
Area Under Curve : 0.7562214